

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 16:30:34 ; Search time 25 Seconds
(without alignments)
1246.770 Million cell updates/sec

Title: US-09-635-501-2

Perfect score: 4291
Sequence: 1 MSSSSWLLLSLVATTAQST.....ISKGENNPGFQNTDDVOTSF 805

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		ID	Description
		Match	Length		
1	1344	31.3	732	1 ACET_HUMAN	P22966 homo sapien
2	1337	31.2	1306	1 ACE_HUMAN	P12821 homo sapien
3	1337	31.2	732	1 ACET_MOUSE	P22967 mus musculu
4	1334	31.1	1312	1 ACE_MOUSE	P09470 mus musculu
5	1312	30.6	1193	1 ACE_CHICK	Q10751 gallus gall
6	1310	30.9	1313	1 ACE_RAT	P47820 rattus norv
7	1283.5	29.9	1310	1 ACE_RABIT	P12822 onyctolagus
8	1283	29.9	737	1 ACET_RABIT	P22968 onyctolagus
9	1090	25.4	615	1 ACE_DROME	Q10715 haematobia
10	1058	24.7	611	1 ACE_HAIE	Q10714 drosophila
11	135	3.1	501	1 YPWA_BACSU	P50848 bacillus su
12	125	2.9	627	1 GLGB_BACSU	P30118 bacillus su
13	123	2.9	986	1 EPTB_STAEP	P30195 staphylococ
14	121	2.8	1034	1 BGAL_BACME	O52847 bacillus me
15	119.5	2.8	3655	1 YAMB_SCHPO	Q10064 schizosacch
16	119.5	2.8	4540	1 DYHC_PARTPE	Q27171 paramecium
17	118	2.7	3911	1 AKN9_HUMAN	Q99996 h a-kinase
18	115.5	2.7	3433	1 UTRN_HUMAN	P46939 homo sapien
19	115	2.7	950	1 Y51L_RICRP	Q92d36 rickettsia
20	114.5	2.7	901	1 P1P_LACLA	P49022 lactococcus
21	114.5	2.7	1225	1 Y309_MYCCE	P47551 mycoplasma
22	113.5	2.6	6669	1 NEBU_HUMAN	P20929 homo sapien
23	112.5	2.6	773	1 CDH_PHACH	Q01738 phanerochae
24	109.5	2.6	1398	1 DNN2_SCHPO	Q9uru2 schizosacch
25	109	2.5	906	1 Y01O_CLOAB	Q97n28 clostridium
26	108.5	2.5	1084	1 XP01_YEAST	P30822 saccharomyc
27	108.5	2.5	2167	1 YGS2_YEAST	P25356 saccharomyc
28	107.5	2.5	1689	1 RPAL_SCHPO	P15398 schizosacch
29	107.5	2.5	3685	1 DMD_HUMAN	P11532 homo sapien
30	107	2.5	736	1 VP4_RPTCP	P26193 porcine rot
31	106.5	2.5	1312	1 RA50_YEAST	P12753 saccharomyc
32	106	2.5	1279	1 BCHH_CHUVI	O50314 chlorobium
33	105.5	2.5	645	1 REP_BUCAI	P57654 buchnera ap

ALIGNMENTS

RESULT 1	ACET_HUMAN	STANDARD:	PRT:	732 AA.
AC	P22966;			
AC	01-AUG-1991 (Rel. 19, Created)			
DT	01-AUG-1991 (Rel. 19, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Angiotensin-converting enzyme, testis-specific isoform precursor (EC 3.4.15.1) (ACE-T) (Dipeptidyl carboxypeptidase I) (Kinase II).			
GN	DCP1 OR DCP OR ACE.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RX	SEQUENCE FROM N.A.			
RX	MEDLINE=90046671; PubMed=2554286;			
RT	Ehlers M.R.W., Fox E.A., Strydom D.J., Riordan J.F.;			
RT	"Molecular cloning of human testicular angiotensin-converting enzyme: the testis isozyme is identical to the C-terminal half of endothelial angiotensin-converting enzyme.";			
RT	Proc. Natl. Acad. Sci. U.S.A. 86:7741-7745(1989).			
RL	[2]			
RN	SEQUENCE FROM N.A.			
RX	MEDLINE=89338720; PubMed=2547653;			
RX	Ra Lattion A.L., Soubrier F., Allegrini J., Hubert C., Corvol P.,			
RA	Alhenc-Gelas F.;			
RT	"The testicular transcript of the angiotensin I-converting enzyme			
RT	encodes for the ancestral, non-duplicated form of the enzyme.";			
RT	FEBS Lett. 252:99-104(1989).			
RL	[3]			
RN	SEQUENCE FROM N.A., AND VARIANTS P-32; G-49 AND S-712.			
RX	MEDLINE=99251580; PubMed=10319862;			
RX	Rieder M.J., Taylor S.L., Clark A.G., Nickerson D.A.;			
RT	"Sequence variation in the human angiotensin converting enzyme.";			
RT	Nat. Genet. 22:59-62(1999).			
RL	[4]			
RN	ZINC-BINDING.			
RP	MEDLINE=91308093; PubMed=1649623;			
RX	Ehlers M.R., Riordan J.F.;			
RT	"Angiotensin-converting enzyme: zinc- and inhibitor-binding			
RT	stoichiometries of the somatic and testis isozymes.";			
RL	Biochemistry 30:7118-7126(1991).			
CC	-1- FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF			
CC	THE TERMINAL HIS-LEU, THIS RESULTS IN AN INCREASE OF THE			
CC	VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.			
CC	-1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,			
CC	oligopeptide-[I-xaa-Xbb, when Xaa is not pro, and Xbb is neither			
CC	Asp nor Glu. Converts angiotensin I to angiotensin II.			
CC	-1- COFACTOR: BINDS 1 ZINC ION.			
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.			
CC	-1- ALTERNATIVE PRODUCTS: THE TESTICULAR ANGIOTENSIN-CONVERTING ENZYME			
CC	IS TRANSCRIBED FROM THE SAME GENE AS THE SOMATIC ISOFORM, PROBABLY			
CC	FROM AN ALTERNATIVE START SITE.			
CC	-1- TISSUE SPECIFICITY: SPERMATOCYTES, ADULT TESTIS.			
CC	-1- INDUCTION: EXPRESSION IS THOUGHT TO BE SUBJECT TO HORMONAL			

REGULATION BY ANDROGENS.
-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M26657; AAA60611.1; -;
DR EMBL; X16295; CAA34362.1; -;
DR EMBL; AF118569; AAD28561.1; -;
DR PIR; S05238; S05238;
DR PIR; A33979; A33979;
DR MEROPS; M02.004; -;
DR MIM; 106180; -;
DR InterPro; IPR001548; Peptidase_M2.
DR InterPro; IPR000130; Zn_MTPeptide.
DR Pfam; PF01401; Peptidase_M2_1.
DR PRINTS; PR00791; PEPTIDPASEA.
DR ProDom; PD004184; Peptidase_M2_1.
DR PROSITE; PS00142; ZINC_PROPEASE; 1.
KW Hydrolase; Metalloprotease; Carboxypeptidase; zinc; Dipeptidase;
KW Glycoprotein; Transmembrane; Testis; Signal; Alternative splicing;
KW Polymorphism.
FT SIGNAL 1 31 ANGIOTENSIN-CONVERTING ENZYME,
FT CHAIN 32 732 TESTIS-SPECIFIC ISOFORM.
FT DOMAIN 32 684 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 685 701 POTENTIAL.
FT DOMAIN 702 732 CYTOPLASMIC (POTENTIAL).
FT METAL 414 414 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 415 415 BY SIMILARITY.
FT METAL 418 418 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 442 442 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 457 457 POTENTIAL.
FT CARBOHYD 103 103 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 121 121 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 368 368 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 617 617 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARIANT 32 32 S -> P.
FT VARIANT 49 49 /FTID=VAR_011710.
FT VARIANT 712 712 /FTID=VAR_011711.
FT VARIANT 712 712 R -> S.
FT SEQUENCE 732 AA; 83330 MW; 80B0D19CFA642313 CRC64;
Query Match 31.3%; Score 1344; DB 1; Length 732;
Best Local Similarity 41.8%; Pred. No. 7.4e-86;
Matches 259; Conservative 119; Mismatches 204; Indels 38; Gaps 10;
QY 15 TAAQS-----TIEQAQTFDKFHEADELFYQSSLASWYNTNITEE-----NVQNM 62
DB 61 TSAQSNLYTDEAEASKFVEEYDRTSQVWVYAEANWYNTNITTSKILLQKNMQIA 120
QY 63 NNAGDKWAFLEKQSTLAOMYPOLEONTLVKQLQALQOQSSVLSDEKSKRLNTILT 122
DB 121 NHT-----LKYQARKFDVNLQNTTIKRIKKVDLRAALPAQLEEYNNKILLD 172
QY 123 MSTYSTGKVCNPDNPQECILLPEGLNEITMANSLDYNERLWAEWSRVSQOLRPLYEE 182
DB 173 METTYSVATVCHPNG--SCLOQEPDTNWNATSRKYEDLLWAEGRDRAGRAILOFYFK 230
QY 183 YVVLKEMARAHYEDYGDYWRGDYEVNGVDGYDSRGOLIEVHTFEIKPLYLEHLA 242
DB 231 YVELINQAARLNGYVDAGDSWRSMYETPSLE-----QDLERLFQELQPLYLNLHA 280
QY 243 YVRACLNNAY-PSYISPIGCLPAHLGLCDMWGRFTWNLVSLTVPFQKPNIDVDMVDQA 301

DB 281 YVRRALHRHYGAQHINLEGPAPAHLLGNMWAQTSNIYDLVPPFPAPSMDTTEAMLKQG 340
QY 302 WDAQRIKFEAKKEFVSGLPNMTQGEWNSMLTDPGNVQKAVCHPTAWDLGKG-DFRIILM 360
DB 341 WTPRMFEKADDFTSGLLLPVPPEFWNKSMLKPTDGRVNVCHASAWDPYNGKDFRIKQ 400
QY 361 CTKVTMDDFTAHEMHGHIQYDMAYAAQPLFLRNGANEGHEAVGEIMLSAATPKHLKS 420
DB 401 CTTVNLEDLVAAHEMHGHIQYFMQYKDLPVALREGANPGFHEAIGDVLALSVSTPKHLHS 460
QY 421 IGLLSPDFQEDNETEINFLIKQALITVGTLPFTYMLEKWEWYFKGEIPKDDMKKWEW 480
DB 461 LNLSSSEGGSD-EHDINFLMKMALDKIAFPFVYLDVQWRVDFGSGTITKENTNQEWSL 519
QY 481 KREIVGVVEPVPHDETCDPASLFHVSNDYSFIRYTRTLYQFOFQALCOAAKHGEGPLH 540
DB 520 RLKYQGLCPVPRTQGDGDFGAKFHIPSSVYIRYFVSFIQFQHEALCOAAHGTPGLH 579
QY 541 KCDISNSTEAGQKLFNMLRLGKSEPTWTLAENVGAKNMVRLNLYFEPFLTWLKDQNK 600
DB 580 KCDIYQSKAQRLATANKLGFSRPMPPEAMQLITGPNMSASAMLSYFKPLDMLRTENE 639
QY 601 --NSFVGH-STDWSPYADQS 617
DB 640 LHGEKLGMPQYNTWTPNSARS 659
RESULT 2
ACE_HUMAN STANDARD; PRT; 1306 AA.
ID AC_P12821;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Angiotensin-converting enzyme, somatic isoform precursor (BC 3.4.15.1)
DE (ACE) (Dipeptidyl carboxypeptidase I) (Kininase II) (CD143 antigen).
GN DCPI OR DCP OR ACE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89071703; PubMed=2849100;
RA Soubrier F., Alhenc-Gelas F., Hubert C., Allegrini J., John M.,
Tregear G., Corbol P.;
RT "Two putative active centers in human angiotensin I-converting enzyme
revealed by molecular cloning."
RL Proc. Natl. Acad. Sci. U.S.A. 85:9386-9390(1988).
[2]
RN SEQUENCE FROM N.A., AND VARIANTS S-261; W-561 AND S-1286.
RX MEDLINE=9251580; PubMed=10319862;
RA Rieder M.J., Taylor S.L., Clark A.G., Nickerson D.A.;
RT "Sequence variation in the human angiotensin converting enzyme."
RL Nat. Genet. 22:59-62(1999).
[3]
RN PARTIAL SEQUENCE OF 30-46.
RC TISSUE=Lung;
RX MEDLINE=90110025; PubMed=2558109;
RA Takeuchi K., Shimizu T., Ohishi N., Seyama Y., Takaku F.;
RT "Purification of human lung angiotensin-converting enzyme by high-
performance liquid chromatography: properties and N-terminal amino
acid sequence."
RL J. Biochem. 106:442-445(1989).
[4]
RN ZINC-BINDING.
RX MEDLINE=91308093; PubMed=1649623;
RA Ehlers M.R., Riordan J.F.;
RT "Angiotensin-converting enzyme: zinc- and inhibitor-binding
stoichiometries of the somatic and testis isozymes."
RL Biochemistry 30:7118-7126(1991).

CC -!- FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF
 CC THE TERMINAL HIS-LEU, THIS RESULTS IN AN INCREASE OF THE
 CC VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.
 CC -!- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
 CC oligopeptide-|-xaa-Abb, when xaa is not pro, and xbb is neither
 CC Asp nor Glu. Converts angiotensin I to angiotensin II.
 CC -!- COFACTOR: BINDS 2 ZINC IONS.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- ALTERNATIVE PRODUCTS: THE TESTICULAR ANGIOTENSIN-CONVERTING ENZYME
 CC IS TRANSCRIBED FROM THE SAME GENE AS THE SOMATIC ISOFORM, PROBABLY
 CC FROM AN ALTERNATIVE START SITE.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.
 CC -!- DATABASE: NAME-PROW; NOTE-CD guide CD143 entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd143.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: J04144; AA51684.1; -.
 CC EMBL: AF118569; AAD28560.1; -.
 CC PIR: A31759; A31759.
 CC PIR: PQ0004; PQ0004.
 CC MIM: 106180; -.
 CC InterPro: IPR001548; Peptidase_M2.
 CC InterPro: IPR000130; Zn_MTPeptidase.
 CC Pfam: PF01401; Peptidase_M2; 2.
 CC PRINTS: PR00791; PEPTIDPASEA.
 CC ProDom: PD0041184; Peptidase_M2; 2.
 CC PROSITE: PS00142; ZINC_PROTEASE; 2.
 CC Hydrolase: Metalloprotease; Carboxypeptidase; Zinc; Dipeptidase;
 CC Glycoprotein; Transmembrane; Repeat; Signal; Alternative splicing;
 CC Polymorphism.
 KW SIGNAL 1 29
 KW CHAIN 30 1306
 FT FT ANGIOTENSIN-CONVERTING ENZYME, SOMATIC
 FT ISOFORM.
 FT DOMAIN 30 1259
 FT TRANSMEM 1260 1276
 FT DOMAIN 1277 1306
 FT REPEAT 227 583
 FT REPEAT 825 1181
 FT METAL 330 390
 FT ACT_SITE 391 391
 FT METAL 394 394
 FT METAL 988 988
 FT ACT_SITE 989 989
 FT METAL 992 992
 FT CARBOHYD 38 38
 FT CARBOHYD 54 54
 FT CARBOHYD 74 74
 FT CARBOHYD 79 79
 FT CARBOHYD 111 111
 FT CARBOHYD 146 146
 FT CARBOHYD 160 160
 FT CARBOHYD 272 272
 FT CARBOHYD 318 318
 FT CARBOHYD 445 445
 FT CARBOHYD 509 509
 FT CARBOHYD 677 677
 FT CARBOHYD 695 695
 FT CARBOHYD 714 714
 FT CARBOHYD 760 760
 FT CARBOHYD 942 942
 FT CARBOHYD 1191 1191
 FT VARIANT 261 261
 FT VARIANT 561 561
 FT VARIANT 1286 1286
 FT R -> W.
 FT /FTid=VAR_011708.
 FT R -> S.
 CC -----
 CC -!- FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF
 CC THE TERMINAL HIS-LEU, THIS RESULTS IN AN INCREASE OF THE
 CC VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.
 CC -!- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
 CC oligopeptide-|-xaa-Abb, when xaa is not pro, and xbb is neither
 CC Asp nor Glu. Converts angiotensin I to angiotensin II.
 CC -!- COFACTOR: BINDS 2 ZINC IONS.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- ALTERNATIVE PRODUCTS: THE TESTICULAR ANGIOTENSIN-CONVERTING ENZYME
 CC IS TRANSCRIBED FROM THE SAME GENE AS THE SOMATIC ISOFORM, PROBABLY
 CC FROM AN ALTERNATIVE START SITE.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.
 CC -!- DATABASE: NAME-PROW; NOTE-CD guide CD143 entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd143.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: J04144; AA51684.1; -.
 CC EMBL: AF118569; AAD28560.1; -.
 CC PIR: A31759; A31759.
 CC PIR: PQ0004; PQ0004.
 CC MIM: 106180; -.
 CC InterPro: IPR001548; Peptidase_M2.
 CC InterPro: IPR000130; Zn_MTPeptidase.
 CC Pfam: PF01401; Peptidase_M2; 2.
 CC PRINTS: PR00791; PEPTIDPASEA.
 CC ProDom: PD0041184; Peptidase_M2; 2.
 CC PROSITE: PS00142; ZINC_PROTEASE; 2.
 CC Hydrolase: Metalloprotease; Carboxypeptidase; Zinc; Dipeptidase;
 CC Glycoprotein; Transmembrane; Repeat; Signal; Alternative splicing;
 CC Polymorphism.
 KW SIGNAL 1 29
 KW CHAIN 30 1306
 FT FT ANGIOTENSIN-CONVERTING ENZYME, SOMATIC
 FT ISOFORM.
 FT DOMAIN 30 1259
 FT TRANSMEM 1260 1276
 FT DOMAIN 1277 1306
 FT REPEAT 227 583
 FT REPEAT 825 1181
 FT METAL 330 390
 FT ACT_SITE 391 391
 FT METAL 394 394
 FT METAL 988 988
 FT ACT_SITE 989 989
 FT METAL 992 992
 FT CARBOHYD 38 38
 FT CARBOHYD 54 54
 FT CARBOHYD 74 74
 FT CARBOHYD 79 79
 FT CARBOHYD 111 111
 FT CARBOHYD 146 146
 FT CARBOHYD 160 160
 FT CARBOHYD 272 272
 FT CARBOHYD 318 318
 FT CARBOHYD 445 445
 FT CARBOHYD 509 509
 FT CARBOHYD 677 677
 FT CARBOHYD 695 695
 FT CARBOHYD 714 714
 FT CARBOHYD 760 760
 FT CARBOHYD 942 942
 FT CARBOHYD 1191 1191
 FT VARIANT 261 261
 FT VARIANT 561 561
 FT VARIANT 1286 1286
 FT R -> W.
 FT /FTid=VAR_011708.
 FT R -> S.

FT CONFLICT 35 35 /FTid=VAR_011709.
 FT CONFLICT 42 42 Q -> E (IN REF. 2).
 FT CONFLICT 42 42 D -> R (IN REF. 2).
 SQ SEQUENCE 1306 AA; 149714 MW; 1B33BCA7301A26AA CRC64;
 Query Match 31.2%; Score 1337; DB 1; Length 1306;
 Best Local Similarity 41.7%; Pred. No. 5.2e-85;
 Matches 255; Conservative 118; Mismatches 204; Indels 34; Gaps 9;
 QY 20 TIEQAKTFLDKFHEAEDLFYQSSLASNNYNTNITEE-----NYQNMMNAGDKWSA 71
 DB 644 TDEAEASKEVEYDTSQVWVNEAEANNYNNTTETSKILLQKNQIAHNT----- 697
 QY 72 FLKEGSTLAQMYPLQEIQLNVKLOLQALQONGSVLSEDSKRLNTILNTMTSIYTGK 131
 DB 698 --LKYGTQARKFDVNVQLQNTIKRIKKVQDLERAALPAQLEEEYNKILLDMETYSVAT 755
 QY 132 VCNPDNPOECILLEPLGMEINMANSIDYNERLWAMESRSEVGKOLRPLYEYVVLKNEMA 191
 DB 756 VCHPNG--SCLOLEPDLTNVATSKYEDLLWAMEGWRDKAGRAILOFPYKYVELINQAA 813
 QY 192 RANHYEDYGDYWRGDIYGVNGVDYDSRGQIEDVDEHTFEIEKPLYEHLHAYVRAKLMA 251
 DB 814 RLNGYVDAGDSWRSWYETPSLE-----QDLERLFQELQPLYLNLHAYVRRALHRH 863
 QY 252 Y-PSVISPIGCLPAHLGLGMGRFWTNLYSLTVPGOKPNTDVTDMYDQANDAQRIPE 310
 DB 864 YGAQHINLEGPPIPAHLGLGNMWAQTSNIYDLVVPFSPASMDTTEAMLKQGTWPRMFE 923
 QY 311 AEKFFVSVGLPNTQGFWENSMKLTDPGNVOKAVCAHPTAWDLGK--DFRILMCTKVTWDDF 369
 DB 924 ADDFTSLGLLPVPPPEFWNKSMLKPTDGRVVCVCHASAWDFYNGKDFRIKQCTTVNLEDL 983
 QY 370 LTAHHENGHIQDYMAAAQPFLLRNGANEGHEAVGEIMSLAATPKHLKIGLLSPDFQ 429
 DB 984 VVAHHENGHIQDYFYQYKDLPLVALREGANPGFHEAIGDVLASVSTPKHLHSLNLSSEGG 1043
 QY 430 EDNTEINFLLKQALTIIVGTLPTFMLEKRWVMVFKGEIPKDKMMKWKWEKRIVGVE 489
 DB 1044 SD-EHDINFMLKALDKIAFIPFSLVDQWRVRVDSITKENTNQENWSUQLYQGLCP 1102
 QY 490 PVPHEITYCDPASLFHVSNDYSFIYRTLYQFQFALCQAAKHEGHLKCDISNSTE 549
 DB 1103 PVPRTQGDGDFGAKFHIPSSVPYIRYFVSFIQFQFHEALCQAAAGHTGPHKCDIYQSK 1162
 QY 550 AGQKLFNMLRGLKSEPTWLALENVVGAKNMVRPLLNYFEPLFTWLKDQNK--NSFVGW- 606
 DB 1163 AGQRLATAMKLGFRPWPPEAMQLITGQPNMSASAMLSYFRKPLLDWURTEHELHGERLW 1222
 QY 607 STDWSPVADQS 617
 DB 1223 QYNWTPNSARS 1233
 RESULT 3
 ACET_MOUSE
 ID ACET_MOUSE STANDARD; PRT; 732 AA.
 AC P22967;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Angiotensin-converting enzyme, testis-specific isoform precursor
 DE (EC 3.4.15.1) (ACE-T) (Dipeptidyl carboxypeptidase 1) (Kininase II).
 GN DCP1 OR ACE.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90318396; PubMed=2164636;
 RA Howard T.E., Shai S.-Y., Langford K.G., Martin B.M., Bernstein K.E.;
 RT "Transcription of testicular angiotensin-converting enzyme (ACE) is

RT initiated within the 12th intron of the somatic ACE gene. *;

RL Mol. Cell. Biol. 10:4294-4302(1990).

CC -!- FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF

CC THE TERMINAL HIS-LEU, THIS RESULTS IN AN INCREASE OF THE

CC VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.

CC -!- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,

CC oligopeptide -|-xaa-xbb, when xaa is not pro, and xbb is neither

CC Asp nor Glu. Converts angiotensin I to angiotensin II.

CC -!- COFACTOR: BINDS 1 ZINC ION

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- ALTERNATIVE PRODUCTS: TRANSCRIPTION OF TESTICULAR ANGIOTENSIN-

CC CONVERTING ENZYME (ACE) IS INITIATED WITHIN THE 12TH INTRON OF

CC THE SOMATIC ACE GENE.

CC -!- TISSUE SPECIFICITY: SPERMATOCYTES, ADULT TESTIS.

CC -!- INDUCTION: EXPRESSION IS THOUGHT TO BE SUBJECT TO HORMONAL

CC REGULATION BY ANDROGENS.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; M55333; AAA37149.1; -;

DR EMBL; M61094; AAA37150.1; -;

DR PIR; A35655; A35655.

DR MEROPS; M02.004; -;

DR MGD; MGI:87874; Ace.

DR InterPro; IPR001548; Peptidase_M2.

DR InterPro; IPR000130; Zn_MTPeptidse.

DR Pfam; PF01401; Peptidase_M2.1.

DR PRINTS; PR00791; PEPTIDPASA.

DR ProDom; PD004184; Peptidase_M2; 1.

DR PROSITE; PS00142; ZINC_PROTEASE; 1.

DR Hydrolase; Metalloprotease; Carboxypeptidase; Zinc; Dipeptidase;

KW Glycoprotein; Transmembrane; Testis; Signal; Alternative splicing.

FT SIGNAL 1 31

FT CHAIN 32 732

FT DOMAIN 32 684

FT TRANSMEM 685 701

FT DOMAIN 702 732

FT METAL 413 414

FT ACT_SITE 414 414

FT METAL 417 417

FT METAL 441 441

FT CARBOHYD 102 102

FT CARBOHYD 120 120

FT CARBOHYD 139 139

FT CARBOHYD 185 185

FT CARBOHYD 367 367

FT CARBOHYD 616 616

SQ SEQUENCE 732 AA; 84047 MW; 16C817E7FBD09BD9 CRC64;

Query Match 31.1%; Score 1334; DB 1; Length 732;

Best Local Similarity 42.6%; Pred. No. 3.7e-85;

Matches 255; Conservative 112; Mismatches 213; Indels 18; Gaps 7;

QY 20 TIEQAKTFLDKFNHBAEDLFYQSSIASWYNTNTNTEENQNNAGDKWAFLEKQSTL 79

DB 69 TDEAKADREVEEDYDRAQVLLNEYAEANQYNTNITIEGSKILLEKSTEVSHTLKYGR 128

QY 80 AQMPYLOETONLTVKLQALQONGSSVLSDEKSKRLNTILNTMTSYTGVCVNDPQ 139

DB 129 AKTFDFSNFNSQSIKKIIXKLQNLDRVLPPKPELYNQILLDMETYSLSLNTG-- 186

QY 140 ECLLEPGLNEIMANSLDNLWAWESRSEVKGOLRPLYEYVVLKNEMARAHYEDY 199

DB 187 TCMPLDPLTNMATSKYKYEELLWAKWSRDKVGRAILPPFPKYVFEFSNKIAKNGYTD 246

QY 200 GDVWRGDEYVNGVDGYDSRGQLIEDVEHTFEEKPLYEHLHAYVRAKLNAYPS-YISP 258

DB 247 GDSWRLSYESDNLE-----QDLKLYQLQLPLYLNLHAYVRRSLHRHYGSEYINL 296

QY 259 IGCPLPAHLGDMGRFWFTNLVSLYTPFGQPNIDVTDAVDQAWDAQRIEKEKEFPVS 318

DB 297 DGPILPAHLGDMMAQTWSNIYDLVAPPSPAPNIDATEAMIKOGMTPRIPKEDNFTSL 356

QY 319 GLPNMTQGFWENSMLTDPGNVQKAVCHPTAWDLGKG--DFRILMCTKVTMDLTAHEMG 377

DB 357 GLLPVPEFVNKSMLEKPTDGRVWVCHPSANDFYNGKDFRIKQCTSVNMDLVIAHEMG 416

QY 378 HIQYDMAAYAAQPFLLRNGANEHFAVGEIMSLSAAPKHLKSLGILLSPDFQEDNFEIN 437

DB 417 HIQYFMQYKDLPTVTFREGANPGFHEAIGDMALSVSPKHLVSLNLLSTE-GSGYEYDIN 475

QY 438 FLKQALITVGLTFTTYMLEKRWNVFKGIPKDOWNKKWEMKREIVGVVEVPVPHDET 497

DB 476 FLKMAWDKIAFIPFSLIDOWRVRVFDGSGITKENYQEWMSURLKYQGLCPVPPRSQGD 535

QY 498 CDPASLFVSNDSYFIRYTRTYLQFOQALCOAAKHEGLHKDCISNSTEAGKLFNM 557

DB 536 FDPGSKFHVANVPYVRYFVSFIQFQFHEALCRAAGTGLHKCDIYQSKAGKLLADA 595

QY 558 LRLKSEPTWLTALENVGAKNMVRLNLYPELFTWLDQNK--NSFVGW-STDNSP 612

DB 596 MGLGYKWPPEAMKLITGQPNMSASAMMYFKPLTEWLVTENRRHGTGLGWPEYNWAP 653

RESULT 4

ACE_MOUSE STANDARD; PRT; 1312 AA.

AC P09470;

DT 01-MAR-1989 (Rel. 10, Created)

DT 01-FEB-1991 (Rel. 17, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE Angiotensin-converting enzyme, somatic isoform precursor (EC 3.4.15.1)

DE (ACE) (Dipeptidyl carboxypeptidase I) (Kininase II).

GN DCPI OR ACE.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_taxid=10090;

FN [1]

RP MEDLINE=89308599; PubMed=2545691;

RA Bernstein K.E., Martin B.M., Edwards A.S., Bernstein E.A.;

RT "Mouse angiotensin-converting enzyme is a protein composed of two

RT homologous domains.";

RL J. Biol. Chem. 264:11945-11951(1989).

[2]

RP SEQUENCE OF 1-332 FROM N.A., AND PARTIAL SEQUENCE.

RA MEDLINE=88298730; PubMed=2841312;

RA Bernstein K., Martin B.M., Bernstein E.A., Linton J., Striker L.,

RA Striker G.;

RT "The isolation of angiotensin-converting enzyme cDNA.";

RL J. Biol. Chem. 263:11021-11024(1988).

CC -!- FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF

CC THE TERMINAL HIS-LEU, THIS RESULTS IN AN INCREASE OF THE

CC VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.

CC -!- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,

CC oligopeptide -|-xaa-xbb, when xaa is not pro, and xbb is neither

CC Asp nor Glu. Converts angiotensin I to angiotensin II.

CC -!- COFACTOR: BINDS 2 ZINC IONS (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- ALTERNATIVE PRODUCTS: THE TESTICULAR ANGIOTENSIN-CONVERTING ENZYME

CC IS TRANSCRIBED FROM THE SAME GENE AS THE SOMATIC ISOFORM, PROBABLY

CC FROM AN ALTERNATIVE START SITE.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

-----
EMBL: J04946; AAA37147.1; -
EMBL: J04947; AAA37148.1; -
EMBL: J03940; AAA37146.1; -
DR PIR: A29220; A29220.
DR PIR: A34171; A34171.
DR MGD: MGI:87874; Ace.
DR InterPro: IPR001548; Peptidase_M2.
DR InterPro: IPR001548; Zn_MTPeptidase.
DR Pfam: PF01401; Peptidase_M2; 2.
DR PRINTS: PR00791; PEPDIPASEA.
DR ProDom: PD004184; Peptidase_M2; 2.
DR PROSITE: PS00142; ZINC_PROTEASE; 2.
KW Hydrolase; Metalloprotease; Carboxypeptidase; Zinc; Dipeptidase;
KW Glycoprotein; Transmembrane; Repeat; Signal; Alternative splicing.
FT SIGNAL 1 34
FT CHAIN 35 1312
      ANGIOTENSIN-CONVERTING ENZYME, SOMATIC
      ISOFORM.
FT DOMAIN 35 1264
FT TRANSMEM 1265 1281
FT DOMAIN 1282 1312
FT REPEAT 232 588
FT REPEAT 830 1186
FT METAL 395 395
FT ACT_SITE 396 396
FT METAL 399 399
FT METAL 993 993
FT ACT_SITE 994 994
FT METAL 997 997
FT CARBOHYD 59 59
FT CARBOHYD 79 79
FT CARBOHYD 116 116
FT CARBOHYD 151 151
FT CARBOHYD 165 165
FT CARBOHYD 323 323
FT CARBOHYD 514 514
FT CARBOHYD 682 682
FT CARBOHYD 700 700
FT CARBOHYD 719 719
FT CARBOHYD 765 765
FT CARBOHYD 947 947
FT CARBOHYD 1196 1196
FT VARIANT 568 568
      T -> A
SQ SEQUENCE 1312 AA; 150947 MW; 9C13BB0529AD3755 CRC64;

Query Match          31.1%; Score 1334; DB 1; Length 1312;
Best Local Similarity 42.6%; Pred. No. 8.4e-85;
Matches 255; Conservative 112; Mismatches 213; Indels 18; Gaps 7;

QY 20 TIEQAKFLDKENHEADLFVQSSLASWYNTNTNTEENVOMNAGDKWSAFLKEQSTL 79
DB 649 TDEAKRDFVEEDYTAQVLLNEYAEANQWYNTNITIEGSKILLEKSTEVSNHLLKYCTR 708
QY 80 AQMPYQLQETQNTVQLQALQOQSSVLSSEKSKRLNTILNTWSTIYTGKVCNPDNPQ 139
DB 709 AKTFVSNFQNSIIRIKKLNQDLRAVLDPKEEYVQIILDMETVYSLSNICYTG-- 766
QY 140 ECLLEPGLNETMANSLDYNERLWAWESWRSVEGKQLRPLYEYVVLKNEMARAHYEDY 199
DB 767 TCMLEPDLTNMATSRYEELLMAWKSWSKVKGRAILPFPPKYVEFSNKTAKLNGYTD 826
QY 200 GDYWRGDEVGVDGYDSRGLIEDVHTTEETKPLVEHLHAYVRAKLMNAYS-YISF 258
DB 827 GDSWSLSYENLSE-----ODLEKLYQELQPLVNLHAYVRRSLHRHGYSEYINL 876
QY 259 IGCLPAHLGDMWGRFTNLSLTVPPFGKPNIDVTDAMVDQAWDAQRIKFAEKFFVSF 318
DB 877 DGPFAHLGNWAGTNSNIYDLVAPFSPAPNIDATEAMIKQGTWTPRIFKEADNFFTSL 936

```

```

QY 319 GLPNWTCQFWENSMITDGNVQKAVCHPTANDLKGK--DFRILMCTKVYTMDDFLFAHBMG 377
DB 937 GLLPVPPPEFWKNSLMLEKPTDGRVYVCHPSAWDFYNGDKFRKQCTSVNMEDLVIAHBMG 996
QY 378 HIQYDMAYAAQPFLLRNGANGEGFHEAVGEIMSLSAATPKHLKSTGLSLGSPDFQEDNETEIN 437
DB 997 HIQYFMQYKDLPTVTFREGANPGFHEAIGDIMALSVSTPKHLYSLNLSLSTE--GSGYEYDIN 1055
QY 438 FLLKQALTYVGTLPFTYMLEKRWMMVFKEIPKDOMKKMMKREIVGVVPEVPHDETY 497
DB 1056 FLMKALDKIAFIPFSYILIDQWRWRVFDGSITKENYQENWNSRLKYQGLCPPVPRSGD 1115
QY 498 CDRASLFHVSNDYSFIIRYVYTTLYQFQEQEALCOAAKHEGPLHKCDTNSNSTEAGOKLFNM 557
DB 1116 FDPGSKFHVPAWVYVYFVFIQFQHEALCRAAGTGPLHKCDIYQSKKAGKLLADA 1175
QY 558 LRLGKSEPTWLLALENVVGAKNMVRPLLYPEPLFTWLKDKNQK--NSFVGV--STDWSP 612
DB 1176 MKLGYSKPWEAMKLITGQPNKSASAMNYKPLTEWLVYTNRRRGHETLGHPEYNWAP 1233

RESULT 5
ACE_CHICK
ID ACE_CHICK STANDARD; PRT; 1193 AA.
AC Q10751:
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Angiotensin-converting enzyme (EC 3.4.15.1) (Dipeptidyl
DE carboxypeptidase I) (Kininase II) (Fragment).
GN DCP1 OR ACE.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=95110342; PubMed=7811282;
RA Esther C.R., Thomas K.E., Bernstein K.P.;
RT "Chicken lacks the testis specific isozyme of angiotensin converting
RL Biochem. Biophys. Res. Commun. 205:1916-1921(1994).
CC -1- FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF
CC THE TERMINAL HIS-LEU, THIS RESULTS IN AN INCREASE OF THE
CC VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.
CC -1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
CC oligopeptide--|-Xaa-Xbb, when Xaa is not Pro, and Xbb is neither
CC Asp nor Glu. Converts angiotensin I to angiotensin II.
CC -1- COFACTOR: BINDS 2 ZINC IONS.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
```

```

-----
EMBL: L40175; AAA5554.1; -
DR InterPro: IPR001548; Peptidase_M2.
DR InterPro: IPR001548; Zn_MTPeptidase.
DR Pfam: PF01401; Peptidase_M2; 2.
DR ProDom: PD004184; Peptidase_M2; 2.
DR PROSITE: PS00142; ZINC_PROTEASE; 2.
KW Hydrolase; Metalloprotease; Carboxypeptidase; Zinc; Dipeptidase;
KW Glycoprotein; Transmembrane; Repeat.
FT NON_TER 1 1
FT DOMAIN <1 1156 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1157 1173

```


FT	METAL	994	994	ZINC 2 (CATALYTIC) (BY SIMILARITY).
FT	ACT_SITE	995	995	2 (BY SIMILARITY).
FT	METAL	998	998	ZINC 2 (CATALYTIC) (BY SIMILARITY).
FT	CARBOHYD	44	44	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	60	60	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	80	80	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	117	117	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	152	152	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	166	166	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	324	324	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	515	515	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	583	583	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	701	701	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	720	720	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	766	766	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	948	948	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1197	1197	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	VARIANT	207	207	R -> K.
SQ	SEQUENCE	1313 AA;	150907 MW;	8CB5D0015F129591 CRC64;
Query Match 30.5%; Score 1310; DB 1; Length 1313;				
Best Local Similarity 42.0%; Pred. No. 3.9e-83;				
Matches 251; Conservative 115; Mismatches 214; Indels 18; Gaps 7;				
QY	20	TIEBQAKFLDKFNHEADLFYQSLSASWNTNTNITEVQNMNAGDKWSAFLKEOSTL	79	
DB	650	TDEAKANFEVEYDRTAKVWNEAEANWHYNTNITIGSKILLQKNKNSHPTLKYGTV	709	
QY	80	AQMYPLQEIQLNITVKLOALQONGSSVLSSEKSKRLNTILNTMSTIYSTGKVCNPDNPQ	139	
DB	710	AKTFDVSFQNSFTIKRIKKYQNVDRVAVLPNPELEEVYQIILDMETTYSVANVCYTING	767	
QY	140	ECLLEPLGNEIMANSIDNERLMAWESWRSEVGKQLRPLRYEYVWLKNEARAHVEDY	199	
DB	768	TCLSLPDLTNIMATSRKYELLWWRKSRDKVGRALIPFPKPYVDSNKIAKLNGSDA	827	
QY	200	GDYWRGDEVNGVDGYSRQGLIEDVEHTEETIKPLYEHLHAYVRAKLMMAYPS-YISP	258	
DB	828	GDSNRSSYESDDLE-----QDLKLYQELQPLYLNLHAYVRSLSHRHYGSEVNL	877	
QY	259	IGCLPAHLGDMGWRFTNLYSLVTFGQKPNIDVTAMVDQANDQAIKPEAKFFVSU	318	
DB	878	DGPFPALHLLGNWQAQTSNIITDLVAPPSAPSIDTAMIKQGWTPRRIFKEADNFTSL	937	
QY	319	GLPNMTQGFWSNMLTDGNYQKAVCHPTAMDLGK-DFRILMCTKVTMDDFLTAHHEMG	377	
DB	938	GILLPVPEFWNKSMLKPTDGRVNVCHASAMDFYNGKDFRIKQCTSYNMEELVLAHHEMG	997	
QY	378	HIQYDMAYAAQPFLLRNCANEGFHEAVGEIMSLSAATPKHLKSGILSPDQEDNTEIN	437	
DB	998	HIQYFMQYKDLVPVTFREGANPGFHEALGDVLALSVSTPKHLHSLNLSSS-EGSYEHDI	1056	
QY	438	FLKQALITVGLTPTMYLWKRWMMFKGETPKDQWKKWEMKREITGVVPEVPHDET	497	
DB	1057	FLMKALDKIAFIPSLIDQWRVRVDGTSIKENYQEWLSRLKYGGLCPVPRSQGD	1116	
QY	498	CDPASLPHVSNDSFYIRYITRTLYQFQFQALCOAAKHEGLHKCDISNSTAGOKLFNM	557	
DB	1117	FDPGSKFHPANVPYIRYIFSIITQFQFHEALCAAGHTGLYKCDIYQSKAGKLLADA	1176	
QY	558	LRLGKSEPTWALENVVCAKMMVRLNLYFEPLFTWLKQDNK--NSFVGV-STDWSP	612	
DB	1177	MKLGYSKQWPEAMKIITGQPNMSASAIMNYFKPLTEWLVTNRHGHGTLGWPEYTWTP	1234	
RESULT 7				
ACE_RABIT	ACE_RABIT	STANDARD;	PRT;	1310 AA.
ID	ACE_RABIT			
AC	P12822; 002852;			
DT	01-Oct-1989 (Rel. 12, Created)			
DT	01-Nov-1997 (Rel. 35, Last sequence update)			
DT	30-May-2000 (Rel. 39, Last annotation update)			
DE	Angiotensin-converting enzyme, somatic isoform precursor (EC 3.4.15.1)			

DE	(ACE) (Dipeptidyl carboxypeptidase I) (Kininase II).
GN	DCP1 OR ACE.
OS	Oryctolagus cuniculus (Rabbit).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX	NCBI_TaxID=9986;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Lung;
RX	MEDLINE=92178960; PubMed=1311831;
RA	Thekkumkara T.J., Livingston W. III, Kumar R.S., Sen G.C.;
RT	*Use of alternative polyadenylation sites for tissue-specific
RT	transcription of two angiotensin-converting enzyme mRNAs.;
RL	Nucleic Acids Res. 20:683-687(1992).
RN	[2]
RP	REVISIONS.
RA	Sen G.C.;
RL	Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN	[3]
RP	SEQUENCE OF 1-88 FROM N.A.
RC	TISSUE=Liver;
RX	MEDLINE=91139683; PubMed=1847388;
RA	Kumar R.S., Thekkumkara T.J., Sen G.C.;
RT	*The mRNAs encoding the two angiotensin-converting isozymes are
RT	transcribed from the same gene by a tissue-specific choice of
RT	alternative transcription initiation sites.;
RL	J. Biol. Chem. 266:3854-3862(1991).
RN	[4]
RP	SEQUENCE OF 34-55.
RC	TISSUE=Lung;
RX	MEDLINE=84051289; PubMed=6314908;
RA	Iwata K., Blacher R., Soffer R.L., Lai C.Y.;
RT	*Rabbit pulmonary angiotensin-converting enzyme: the NH2-terminal
RT	fragment with enzymatic activity and its formation from the native
RT	enzyme by NH4OH treatment.;
RL	Arch. Biochem. Biophys. 227:188-201(1983).
CC	-1- FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF
CC	THE TERMINAL HIS-LEU. THIS RESULTS IN AN INCREASE OF THE
CC	VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.
CC	-1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
CC	oligopeptide-l-Xaa-Xbb, when Xaa is not Pro, and Xbb is neither
CC	Asp nor Glu. Converts angiotensin I to angiotensin II.
CC	-1- COFACTOR: BINDS 2 ZINC IONS (BY SIMILARITY).
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.
CC	-1- ALTERNATIVE PRODUCTS: THE TESTICULAR ANGIOTENSIN-CONVERTING ENZYME
CC	IS TRANSCRIBED FROM THE SAME GENE AS THE SOMATIC ISOFORM, PROBABLY
CC	FROM AN ALTERNATIVE START SITE.
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; X62551; CAA44428.1;
DR	EMBL; M58579; AAA31151.1; ALT_SEQ.
DR	PIR; A23455; A23455.
DR	PIR; S35484; S35484.
DR	InterPro; IPR001548; Peptidase_M2.
DR	InterPro; IPR00130; Zn_Mtpeptdse.
DR	Pfam; PF01401; Peptidase_M2; 2.
DR	PRINTS; PR00791; PEPDIPASEA.
DR	ProDom; PD004184; Peptidase_M2; 2.
DR	PROSITE; PS00142; ZINC_PROTEASE; 2.
KW	Hydrolase; Metalloprotease; Carboxypeptidase; Zinc; Dipeptidase;
FT	Glycoprotein; Transmembrane; Repeat; Signal; Alternative splicing.
FT	SIGNAL 1 33
FT	CHAIN 34 1310 ANGIOTENSIN-CONVERTING ENZYME, SOMATIC
FT	ISOFORM.
FT	DOMAIN 34 1263 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 1264 1280 POTENTIAL.
FT DOMAIN 1310 1310 CYTOPLASMIC (POTENTIAL).
FT REPEAT 322 588
FT REPEAT 829 1185
FT METAL 395 395
FT ACT_SITE 396 396 ZINC 1 (CATALYTIC) (BY SIMILARITY).
FT METAL 399 399 ZINC 1 (CATALYTIC) (BY SIMILARITY).
FT METAL 992 992 ZINC 2 (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 993 993 ZINC 2 (CATALYTIC) (BY SIMILARITY).
FT METAL 996 996 ZINC 2 (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD 59 59 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 79 79 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 151 151 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 323 323 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 449 449 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 513 513 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 681 681 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 699 699 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 718 718 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 946 946 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1195 1195 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 48 48 E -> N (IN REF. 4).
SQ SEQUENCE 1310 AA; 150405 MW; 04777FAB1798IDEA CRC64;

Query Match 29.9%; Score 1283.5; DB 1; Length 1310;
Best Local Similarity 36.8%; Pred. No. 2.7e-81;
Matches 275; Conservative 135; Mismatches 262; Indels 75; Gaps 17;

QY 7 LLLSLVAVTAAGSTI-----EEOAKTFLDKFNHAEADLFYQSSLASWYNTNI 54
DB 21 LLLLLRPPPAALTDPGLPLGDFADAEAGARLFASSYNSAEQVLFSTRSTAASWAHDNI 80

QY 55 TEENVMNNAGDKWGAFLKQSTLAQMYPLQEI-----QNLIT---VKLQALQALQNGSS 106
DB 81 TAENARQEE-----EALLSQFEAEAGKAKELYDPWQNTDFELRRIGAVRTLGA 135

QY 107 VLSEDKSKRLNTLNTMTSTYTGKVCNPDNPQECILLLEPGNEIMANSLOYNERLWAE 166
DB 136 NLPLAKROOYSLLSNNSQIYTGKVCNPDNPQECILLLEPGNEIMANSLOYNERLWAE 195

QY 167 SWSEVKGOLRPLYEVVVLKMEMARHVEDYGDYWRGDEYVNGVDGYDSRGGLIEDV 226
DB 196 GWINAVGIPLKPLYQETALSNEAYRQDGFSDTGATWRSYDSPTFE-----EDL 245

QY 227 EHTFEIKPLYLHAYVRKALNNAY-PSYISPIGCPALHLLGDMGRETWNLVSLTVP 285
DB 246 ERIYHOLEPLYNLHAYVRRVLRHRRYGDYINLRGPIPAHLGNNMAQSWESIYDMMVVF 305

QY 286 GQPNIDVTDAMVDQAMDQRIKFAEKFFVSVGLPNMTQGTWENSLMDPCNVOKAVCH 345
DB 306 PDKPNLDVTSVMYQGNATHMFRVAEEFTSLGLLPPMPPETFAESMLEKPEDGREVVCH 365

QY 346 PTAWDL-GKGDFFILMCTVTDDELTAHHEGHIOYDMAYAAQPLLRNANGEGFHEAV 404
DB 366 ASAWDFYNRKDFIKOCTQVMDOLSTVHEMGHVQYIYQYKQDPVSLR-ANPGFHEAI 424

QY 405 GEIMSLSAATPKRLKSIGLSPDFQEDNETEINFLKQALITVGLPFTYMLERKRWYVF 464
DB 425 GDVLLSVSTPAHLKIGLLD-HVTNDESDINYLKMALEKIAFLPPCYLYVDQNRWGVF 483

QY 465 KGEIPKQMKMKWEMKREIVGVVEPVPHDEYCDPASLPHVSNDSYFRYTRILYQFO 524
DB 484 SGRTPSRVNFDMWYLRKYQICPPVVRNETHFDAGAKFHPSPVTPYRYFVSFVLOQP 543

QY 525 FOEALQAAKHEGLHKCDISNSTEAGOKLFNMLRGKSEPTWTLALENVVGAKNMVRPL 584
DB 544 FHALCWEAGHQPGLHQCDDIYQSTRAGAKLRAVLQAGCSRPHQEVLLKDMVASDALDAQPL 603

QY 585 LNYFEPLFTWLKQDNKNS--FVGV-STDMSPYAQDSIKVIRLSKLSALG-----DKA 632
DB 604 LDYFQPVQWOLQONERNNGEVLGWPEYQWRPPLPNPNYPEGIDLVTDEABASRFEVEYDRS 663

QY 633 YE--WND-----NEMVLRSSVAYAMROYFLKVKVKNQMLFGEDYRVANLKPRIISFNEF 684

DB 664 FOAVWNEYAEANNYNITNTITTEASKILLQKNQIANHTLTYG-----NWARRFDVSNF 716
QY 685 VTAP-----KNVSD-----IIPRTEVEK 702
DB 717 QNATSKRIIRKKVQDLQRAVLVKELEE 743

RESULT 8
ACET_RABIT STANDARD; PRT; 737 AA.
ID ACET_RABIT
AC P22968;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Angiotensin-converting enzyme, testis-specific isoform precursor
DE (EC 3.4.15.1) (ACE-T) (Dipeptidyl carboxypeptidase I) (Kinase II).
GN DCP1 OR ACE.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEW ZEALAND WHITE; TISSUE=Testis;
RX MEDLINE=89380303; PubMed=2550457;
RA Kumar R.S., Kusari J., Roy S.N., Soffer R.L., Sen G.C.;
RT "Structure of testicular angiotensin-converting enzyme. A segmental
RT mosaic isozyme.";
RL J. Biol. Chem. 264:16754-16758(1989).
RN [2]
RP SEQUENCE OF 1-117 FROM N.A.
RX MEDLINE=91139683; PubMed=1847388;
RA Kumar R.S., Thekkumkara T.J., Sen G.C.;
RT "The mRNAs encoding the two angiotensin-converting isozymes are
RT transcribed from the same gene by a tissue-specific choice of
RT alternative transcription initiation sites.";
RL J. Biol. Chem. 266:3854-3862(1991).
CC -1- FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF
CC THE TERMINAL HIS-LEU, THIS RESULTS IN AN INCREASE OF THE
CC VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.
CC -1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
CC oligopeptide-l-Xaa-Xbb, when Xaa is not Pro, and Xbb is neither
CC Asp nor Glu. Converts angiotensin I to angiotensin II.
CC -1- COFACTOR: BINDS 1 ZINC ION.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS: THE TESTICULAR ANGIOTENSIN-CONVERTING ENZYME
CC IS TRANSCRIBED FROM THE SAME GENE AS THE SOMATIC ISOFORM, PROBABLY
CC FROM AN ALTERNATIVE START SITE.
CC -1- TISSUE SPECIFICITY: SPERMATOCYTES, ADULT TESTIS.
CC -1- INDUCTION: EXPRESSION IS THOUGHT TO BE SUBJECT TO HORMONAL
CC REGULATION BY ANDROGENS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR J05041; AAA31153.1; -;
DR EMBL; M58580; AAA31152.1; -;
DR PIR; A34402; A34402.
DR MEROPS; M02.004; -;
DR InterPro; IPR001548; Peptidase_M2.
DR InterPro; IPR000130; Zn_MTPeptide.
DR Pfam; PF01401; Peptidase_M2; 1.
DR PRINTS; PR00791; PEPTIDASEA.
DR PRODom; PD004184; Peptidase_M2; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Metalloprotease; Carboxypeptidase; Zinc; Dipeptidase;

KW Glycoprotein; Transmembrane; Testis; Signal; Alternative splicing.
 FT SIGNAL 1 32
 FT CHAIN 33 737
 FT DOMAIN 33 590
 FT TRANSMEM 691 707
 FT DOMAIN 708 737
 FT METAL 419 439
 FT ACT_SITE 420 420
 FT METAL 423 423
 FT METAL 447 447
 FT CARBOHYD 108 108
 FT CARBOHYD 126 126
 FT CARBOHYD 145 145
 FT CARBOHYD 373 373
 FT CARBOHYD 622 622
 SQ SEQUENCE 737 AA; 83923 MW; FC43CC7655C3DCA CRC64;

Query Match 29.9%; Score 1283; DB 1; Length 737;
 Best Local Similarity 40.8%; Pred. No. 1.3e-81;
 Matches 250; Conservative 112; Mismatches 213; Indels 38; Gaps 10;

QY 20 TLEQAKTFLDKFHEAEDLYQSSLASWNTNITEE-----NVONNN--AGDKW 69
 Db 75 TDEARSFVEYDSFOAVNWEAEANNTNITTEASKILLQKNQIANHTLYGNW 134
 QY 70 SAFLKEQSTLAQMYPLQIQNLTVKQLQALQNGSSVLSSEKSKRLNTILNTMTSTYST 129
 Db 135 -----ARRDVSFNQATSKRIKKVQDQRLQVLPVKEEYQAILDMETIYSV 184
 QY 130 KVCNPDNPQECLELLEGLNEMANSNDYRLNMAWESWRSEVGKQLRPLYEYVVLKNE 189
 Db 185 ANVCRVDG--SCLQLEPLDTNLMTATSRKYDELLWMTWSRDKVGRAILPPYKVEFTNK 242
 QY 190 MARAHYEDYGDYWRGDEYVNGVDGYDSRGQLIEDVEHTEERIKPLYEHLHAYVRAKLM 249
 Db 243 AARLNGYVDAGDSWRSMTETPLE-----ODLERLFOELQPLYNLHAYVGRALH 292
 QY 250 MAY-PSYISPIGCLPAHLGLDMWGRFTNLSLTPFGQKPNIDVTAMVDQANDQRIIF 308
 Db 293 RHYGAOHINLESPIPAHLLGNMWAQTSNIYDLVAPPSASTMDATAMTKQGTTPRMF 352
 QY 309 KEAEKFFVSVGLPNWGTGFNSMLTDPGNVQKAVCHPTAWDLGK--DFRILMCTKYTMD 367
 Db 353 EADKFFLSGLLPVPEFVFNKSMLEKPTDGRVNVCHASAWDFYNGDKFRICKQCTVWME 412
 QY 368 DFLTAHBMGHIQYDMAAOPFLLRNGAGFHEAEGEIMSLSAATPKHLKSLGILLSPD 427
 Db 413 DLVVVHBMGHIQYMQYKDLPLVALRGANPGFHEAIGDVLALSVPKHLHSLNLSSE 472
 QY 428 FOEDNETEINFLKQALTIIVTGLPFTYMLEKRWVFKGEIPKDWMMKWKWEMKREIVGV 487
 Db 473 -GGGYEHDINFLMKALDKIAFIPFSLYLDVDRWRVFDGSGTITKENYQEWWSLRLKQGL 531
 QY 488 VEPVPHETDYCPASLHVSNDYSFIRYTRTLQFOFQALCAQAEKHEGLPKHCKDSNS 547
 Db 532 CPPAPRSGDDPGAKPHIPSSVPIRIFVFSFIQFHEALCAAGTGLPHTCDIYQS 591
 QY 548 TEAQKLFNMLRGLKSPWTLALNNVYGNKMMNVRPLINYPFLFTWLKQDN--KNSFVG 605
 Db 592 KEACKRLADAMKLGYSKFWPEAMKVIQGNMNASAMNFKPLMDMLLTENGHRHGKLG 651
 QY 606 W-STDNSPYADQS 617
 Db 652 WPOYTWTPNSARS 664

RESULT 9
 ID ACE_DROME STANDARD; PRT; 615 AA.
 AC Q10714; Q27572;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)

30-MAY-2000 (Rel. 39, Last annotation update)
 Angiotensin-converting enzyme precursor (EC 3.4.15.1) (Dipeptidyl
 carboxypeptidase I) (Kininase II).
 ANCE OR RACE.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95293950; PubMed=7775412;
 RA Cornell M.J., Williams T.A., Lamango N.S., Coates D., Corvol P.,
 Soubrier F., Honeisel J., Lehrach H., Isaac R.E.;
 FT Cloning and expression of an evolutionarily conserved single-domain
 FT angiotensin converting enzyme from Drosophila melanogaster.;
 RL J. Biol. Chem. 270:13613-13619(1995).
 RN [2]
 RP REVISIONS.
 RA Cornwell M.J.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CANTON-S;
 RX MEDLINE=96028519; PubMed=7547464;
 RA Tatei K., Cai H., Ip Y.T., Levine M.;
 FT Race: a Drosophila homologue of the angiotensin converting enzyme.;
 RL Mech. Dev. 51:157-168(1995).
 CC -1- FUNCTION: MAY PLAY A ROLE IN THE CONTRACTIONS OF THE HEART, GUT
 CC AND TESTES.
 CC -1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
 CC oligopeptide-[Xaa-Xbb, when Xaa is not Pro, and Xbb is neither
 CC Asp nor Glu. Converts angiotensin I to angiotensin II.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- COFACTOR: BINDS 1 ZINC ION.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN THE AMNIOEROSA DURING GERM BAND
 CC ELONGATION, SHORTENING AND HEART MORPHOGENESIS. EXPRESSED IN
 CC MIDGUT THROUGHOUT EMBRYOGENESIS.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U25344; AAC02171.1; -;
 CC EMBL; U34599; AAC46902.1; -;
 CC MEROPS; M02.003; -;
 CC FlyBase; Fgn0012037; Ance.
 CC InterPro; IPR001548; Peptidase_M2.
 CC InterPro; IPR001330; Zn_MTpeptdse.
 CC Pfam; PF01401; Peptidase_M2; 1.
 CC PRINTS; PR00791; PEPTIDASEA.
 CC ProDom; PD004184; Peptidase_M2; 1.
 CC PROSITE; PS00142; ZINC_PROTEASE; 1.
 CC Hydrolase; Metalloprotease; Carboxypeptidase; Zinc; Dipeptidase;
 KW Glycoprotein; Signal.
 FT SIGNAL 1 17
 FT CHAIN 18 615
 FT METAL 367 367
 FT ACT_SITE 368 368
 FT METAL 371 371
 FT CARBOHYD 53 53
 FT CARBOHYD 196 196
 FT CARBOHYD 311 311
 FT CONFLICT 48 51
 FT CONFLICT 141 141
 FT CONFLICT 293 293
 FT CONFLICT 486 486
 FT CONFLICT 533 533

FT	CONFLICT	547	547	R > A (IN REF. 3).
SQ	SEQUENCE	615 AA:	71025 MW;	3583DD5EAF7F3CFB CRC64;
	Query Match	25.4%;	Score 1090;	DB 1; Length 615;
	Best Local Similarity	36.6%;	Pred. No. 2.7e-68;	
	Matches 224;	Conservative 122;	Mismatches 244;	Indels 22; Gaps
Qy	8	L L S I V A V T A A Q S T I B E Q A T F D K F N H E A D L F Y Q S S L A S W N Y N I T E E N V Q O N M N A C D	67	
Db	8	L L A T V A T Q A L V K E E I Q A K E Y L E N L K E L A K R T N I V E T A A W A Y S G N I D T E K K K N E I S A	67	
Qy	68	K W A S F L T K E O S T L A Q M Y P L O E I O N L T V K L Q L A Q L O O Q S S V L S E D K S K R L N T I L N T M S T I Y	127	
Db	68	E L A K F M K E V A S D T T K F Q N S Y Q S E D L K R F K A L T K L G Y A A L P E D D Y A E L D T L S A M E S N F	127	
Qy	128	S T G K V C N P D P O B C - L I L L E P F L G N I M A N S L D Y N E R L A W E S W R S E V G K O L P R Y E E Y V L	186	
Db	128	A K V K C D Y K D S T K C D L A L D P E I B E V I S K S D R H E E L A Y Y W R E P V D K A G T A V R S O F E R Y V E L	187	
Qy	187	K N E M A R A N H Y E D Y G D Y W R G D Y E V N G Y D Y S R G Q L I E D V H T F F E I R P L Y E H L H A Y R A	246	
Db	188	N T K A A K L N N T S G A E A W L D E Y E ----- D D T F Q O Q L E D I --- F A D I R P L X Q O I H G Y V R F	237	
Qy	247	K L M N A Y - P S Y I S P T G C L P A H L L G M D M G R F W T N L Y S U T V P F G K P N I D V T D A M D V Q A D A Q	305	
Db	238	R L R K H Y G D A V S E T G P I P H L L G N M A Q O S E L A D I V S P F P E K P L V D Y S A E M E K Q A Y T P L	297	
Qy	306	R I F E A E K F F V S V G L P N M T G G F W E N S H L T D P G N V Q K A V C H P T A W D L G K - G D P R I L M C T K V	364	
Db	298	K M F Q M G D D F T S M N L T K L P Q D F W D K S I E K P T D G R L V C H A S A W D P Y L D D D V R I K Q C T R V	357	
Qy	365	T M D D F L T A H E M G H I Q Y D M A Y A A Q P L L R N G A N E G F H E A V G E T M S L S A A T P K H L K S I G L L	424	
Db	358	T Q D Q L F T V H H E L G H I F F L Q Y Q O P F Y R T G A N G P H E A V G D V L S U S T P P K H L E K I G L L	417	
Qy	425	S P D F Q E D N E T E I N F L K Q A L T I V G T L P F T Y M L E K W M Y F K G E I P K D O R M K K W M K R E I	484	
Db	418	K - D Y V R D D E A R I N Q L F T A L D K I V F L P P A F T M D K Y R M S L F R G E V D K A N N C A F W K L R D E Y	476	
Qy	485	V G V V E P P H D E T Y C D P A S L F H V S N D S F I R Y T R T L Y Q F Q F A L C - Q A A K H - - - - - E G P	538	
Db	477	S G I E P P V V R S E K O F D A P A K Y H I S A D V E Y L R Y L V S F T I Q F Y K S A C I K A G O Y D P D N V E L P	536	
Qy	539	L H K D C I G N S T E A G K L P N M L R L G K S E P W T L A L E N N Y V G A K N M N V R P L L N T F E P L F T W L K Q	598	
Db	537	L D N C D I G S A R A G A A F H M L S M G A K P W P D A L E A F N G E R I M S G K A T A E F P E P L R V L W L E A F	596	
Qy	599	N - - K N S F V G W S T	608	
Db	597	N I K N N V H I G W T T	608	

	RESULT	10	
ID	ACE_HAEIE		
ID	ACE_HAEEI	STANDARD;	PRT; 611 AA.
AC	Q10715;		
DT	01-NOV-1997	(Rel. 35,	Created)
DT	01-NOV-1997	(Rel. 35,	Last sequence update)
DT	30-MAY-2000	(Rel. 39,	Last annotation update)
DE	Angiotensin-converting enzyme precursor [EC 3.4.15.1] (Dipeptidyl carboxypeptidase I) (Kininase II).		
GN	ACE.		
OS	<i>Haematobia irritans exigua</i> (Buffalo fly).		
OC	Eukaryota:	Metazoa:	Arthropoda:
OC	Pterygota:	Neoptera:	Endopterygota:
OC	Muscoidea:	Muscidae:	Haematobinae.
OX	NCB1_TaxID=34678;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=96215437;	PubMed=8647080;	
RA	Wijffels G.L., Fitzgerald C., Gough J., Riding G.A., Elvin C., Kemp D.J., Willadsen P.;		
RA	"Cloning and characterisation of angiotensin-converting enzyme from		

[illegible]

QY 424 LSPDFQEDNETEFLLKQALITVGLTPPTMYLERWVFKGEIPKQWMMKMKRE 483
 Db 417 LK-NTVSDNEARINQFLTALDKIVFLPFAFTMDKYRWALFRGQADKSEWCAFWKLE 475
 QY 484 IVGVVPEPHDTCYCDPASLHVSNDSYFIRYRTLYQFOFQALCOAA-----KHG 537
 Db 476 YSGIEPPVVRTEKDPDAKHYVSADVEYLYRYSFIQFQYKSACTAGAYVPNOT 535
 QY 538 PLHKCDISNSTEAGKLFENMLRGLKSEPTLALNNVCAKMMNVRPLNLYFEPLFTWLK 596
 Db 536 PLDNCDIYSGKAGKLFENMLSLGSKPWPDALEAFNGERTWGTAKIAYEPLRVWLE 594

RESULT 11

ID YPWA_BACSU STANDARD; PRT; 501 AA.
 AC P50848;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 58.2 kDa protein in KDGT-XPT intergenic region.
 GN YPWA.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxId=1423;
 RN [1]
 RC SEQUENCE FROM N.A.
 RX STRAIN=168 / MARBURG;
 RX MEDLINE=96349105; PubMed=8760912;
 RA SOROKIN A.V., AZEVEDO V., ZUMSTEIN E., GALLERON N., EHRLICH S.D.,
 RA SERROR P.;
 RT "Sequence analysis of the Bacillus subtilis chromosome region between
 the sera and kdg loci cloned in a yeast artificial chromosome.";
 RL Microbiology 142:2005-2016(1996).
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M32 (ZINC
 METALLOPROTEASE).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L47838; AAB38482.1; -;
 DR EMBL; L77246; AAA96610.1; -;
 DR EMBL; Z99115; CAB1425.1; -;
 DR MEROPS; M32.001; -;
 DR Subtilist; BG11458; YPWA.
 DR InterPro; IPR001333; Peptidase_M32.
 DR InterPro; IPR000130; Zn_MTPeptidse.
 DR Pfam; PF02074; Peptidase_M32; 1.
 DR PRINTS; PR00998; CR00XPTASET.
 DR PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
 KW Hypothetical proteome; Hydrolase; Metalloprotease; Zinc;
 KW Complete proteome.
 FT METAL 265 265 ZINC (CATALYTIC) (POTENTIAL).
 FT ACT_SITE 266 266 POTENTIAL.
 FT METAL 269 269 ZINC (CATALYTIC) (POTENTIAL).
 SQ SEQUENCE 501 AA; 58174 MW; A7489BBAFEFA38F82 CRC64;

Query Match 3.1%; Score 135; DB 1; Length 501;
 Best Local Similarity 20.9%; Pred. No. 0.056;
 Matches 98; Conservative 65; Mismatches 164; Indels 142; Gaps 26;

QY 180 YEEYVWL-KNEMA--RANHVEDYG-----DYWRGDYEVNGVDGYDS 218
 Db 103 YKEYVILCSKAETAEWEAKGSDPSFLSPYLEQLIEPKRFTYV-----GVQ-- 150
 QY 219 RGQLIEDVEHTFEIKPLYEHLHAYVAKLMNAYPSIPIGCLPAHLGLDMGRFTWNL 278

Db 151 -----EHPYDALLDLFE---PGVTVKYLD-----QLFAEL 177
 QY 279 YSLTVPF-----GQPNIDVTDAMVDQANDQARIPEAKEFFVSVGLPNMTQGFWENS 331
 Db 178 KEALIPLVKQVTSAGNKP-----DTSFITKAPFKPKQKLSLYFLQELGYDF----- 224
 QY 332 MLTDGNGVOACVCHPTAWDLKGGDFRILMCTKVTMMDDFLTA---HHEMGIQYD--MAY 385
 Db 225 ---DGGRLDEV-HPFATTLNRGDRV--TTRYDEKDFRTAIFGTHECGHAIYEQNI 278
 QY 386 AAQPFLLRNGANEGFHEA-----VGE-----IMLSAATPKHLKKSIGLLSPDF- 428
 Db 279 ALSGTNLSDGASMGIHESQSILFVENFGRNKHFWTPYKKIQEASPVQFKDISL--DDFV 336
 QY 429 QEDNETEINFLKQA--LTIVGLTPPTMYLEKRWVNFKEIPKQWMMKMKREIVG 486
 Db 337 RAINEKSPSIRVEADELTYPHLIIIRYEIEK---AIFSNEVSVEDLPSLWNNKYQDYLG 393
 QY 487 VPEVPDHETDCYPASLHVSNDSY-FIRYRTLYQFOFQALCO-----AAKHGGL 539
 Db 394 I---TPQTDAGILQDVHWAGGDFGYPFVALGYMTAAQLKQMLDLPFDLALERGEF 450
 QY 540 HKCDISNSTEAGKLFENMLRL-GKSEPTLALNNVCAKMMNVRPLNLY 587
 Db 451 HPIK-----OWLTEKVHIHGKRRKPLDIKDATG-EELNVRYLIDY 490

RESULT 12

ID GLGB_BACSU STANDARD; PRT; 627 AA.
 AC P391118;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching
 enzyme).
 DE GLGB.
 GN GLGB.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxId=1423;
 RN [1]
 RC SEQUENCE FROM N.A.
 RX STRAIN=168;
 RX MEDLINE=94195107; PubMed=8145641;
 RA Kiel J.A.K.W., Boels J.M., Beldman G., Venema G.;
 RT "Glycogen in Bacillus subtilis: molecular characterization of an
 operon encoding enzymes involved in glycogen biosynthesis and
 degradation.";
 RT degradation.";
 RL Mol. Microbiol. 11:203-218(1994).
 RN [2]
 RC SEQUENCE FROM N.A.
 RX MEDLINE=98048467; PubMed=9387221;
 RA Lapidus A., Galleron N., Sorokin A., Ehrlich S.D.;
 RT "Sequencing and functional annotation of the Bacillus subtilis genes
 in the 200 kb rnb-dnaB region.";
 RT Microbiology 143:3431-3441(1997).
 RL Microbiology 143:3431-3441(1997).
 CC -!- FUNCTION: CATALYSES THE FORMATION OF THE ALPHA-1,6-GLUCOSIDIC
 LINKAGES IN GLYCOGEN BY SCLSSION OF A 1,4-ALPHA-LINKED
 OLIGOSACCHARIDE FROM GROWING ALPHA-1,4-GLUCAN CHAINS AND THE
 SUBSEQUENT TRANSFER OF THE OLIGOSACCHARIDE TO OTHER PARTS OF
 ALPHA-1,4-GLUCAN CHAINS.
 CC -!- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of
 glycogen.
 CC -!- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
 CC -!- INDUCTION: EXPRESSED EXCLUSIVELY ON MEDIA CONTAINING CARBON
 SOURCES THAT ALLOW EFFICIENT SPOULATION.
 CC -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; 225795; CAA81040.1; -
DR EMBL; AF008220; AAC00214.1; -
DR EMBL; Z99119; CAB15076.1; -
DR PIR; S36624; S36624.
DR Subtilisin; BGL10907; glgB.
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR004193; isomylase.N.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isomylase.N; 1.
KW Glycogen biosynthesis; Transferase; Glycosyltransferase;
KW Complete proteome.
FT ACT_SITE 309 BY SIMILARITY.
FT ACT_SITE 352 BY SIMILARITY.
FT ACT_SITE 420 BY SIMILARITY.
SQ SEQUENCE 627 AA; 73665 MW; 64B0A553B6767BA CRC64;

Query Match 2.9%; Score 125; DB 1; Length 627;
Best Local Similarity 18.6%; Pred. No. 0.38; 225; Indels 264; Gaps 37;
Matches 131; Conservative 86; Mismatches 86;

QY 49 NYNTNITEENVONMAGDKSAFL-----KEQSTLAQMYPLQETQNTVRLQALQON 103
DB 56 DNSWSGEEHVMRYNDNCIWTFLPIGKER-----YKYEIVN-----N 97
QY 104 GSSVLSSEKSKRLNLTMTSTYTGKVCNPDNPQECILLLEPGLNETMAN--SLDYNER 161
DB 98 GRIKLKADP-----YALYS-----EVRNPTASLTIDLE 125
QY 162 LAWESRSREVGKQLRPLEYEVVVLKNEARAHYEDYDGRGVEVNGVDYDSRGQ 221
DB 126 GYSWQDQKQKQAKTKYEKPVFI-----YELHLSWK-----KHSDGRHSYKE 171
QY 222 LIEDV-----EHTFE--EIKPLYEHLHAYVRAKLMNAYPSYISPGCLPAHLGLDMWGRF 274
DB 172 LSQTLIPYIKKHGFTHIELLPVE--HPYDRS-----WGYD 205
QY 275 WTNLYSLVFPQCKNIDVTDAMVQA-----WDAQRIFKEAEKFFSVSLPNNMT 324
DB 206 GTGYSPSTRFG--PPHDLM-KFVDECHQONTGVILWVPGHCKADAGLNFDEGP--- 259
QY 325 QGFWENSLTDPGNYQKAVCHPTAMDGLKGRPRILMCKTKVTMDDELTAH-----HEMGHIQ 380
DB 260 --LYEYKEERDREN-----NLWGTANFDL--GKPEVHSFLSNALYWAEFYH- 303
QY 381 YDMAYAAQPFLLRNGANEGFH-EAYGEINSLSAATPKHLKSTGLLSLPDQEDNETEINFL 439
DB 304 -----DGRFVDAVANILYWPNDQDERH-----TNPYAVDFL 333
QY 440 LQAQTLVGTLPFTYML--EKRWVMVFKEIPKD-----QWKKWEMKREIVGVVEPV 491
DB 334 KKLNTQMRRAYPHVMIAEDSTENFQVTVGAVEEGGLGPHYKWMGW---MNDVLKYMETP 390
QY 492 PHDETCDPASLHVSNDSYFIRYTRTLTYQFOQFALCQAAKHGEPHLKCDISNTRAG 551
DB 391 PEERRHC-----HQLISFSL--YAFSEHFVLPFSHDEVV-----G 425
QY 552 QKLFNMLRGKSEPWTALENVVGAKNMVRPLNY-----PEPLFTWLK 596
DB 426 KK--SLLNKMPGDYW-----OKFAQYRLLLGYMTVHPGKKLIFMGSEFAQFDW-K 473
QY 597 DONKNSFCVQSTWSPYADQSI-----KYRISLKSALGDKAYEND---NEMYL 642
DB 474 DTEQ---LDWFDSFPMHQKASVFTQDLRLFYQKSKILYEHDRHQASEMIDVHDEQSI 530
QY 643 FRSSVAYAMROYFLKVKNQMLIFGEHEDVRANLKPRISENFVFTAP 688
DB 531 F-SFIYQKHK-----GEALVICNFTPVVYHOYDVGVP 563

RESULT 13
EPIB_STAEP
ID EPIB_STAEP STANDARD; PRT; 986 AA.
AC P30195;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Epidermin biosynthesis protein epib.
GN EPIB
OS Staphylococcus epidermidis.
OG Plasmid ptu 32.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RU 3298 / DSM 3095;
RX MEDLINE=92155237; PubMed=1740156;
RA Schnell N., Engelke G., Augustin J., Rosenstein R., Ungermann V.,
RA Goetz F., Entlian K.-D.;
RT *Analysis of genes involved in the biosynthesis of lantibiotic
epidermin*;
RL Eur. J. Biochem. 204:57-68(1992).
CC -!- FUNCTION: INVOLVED IN THE POSTTRANSLATIONAL MODIFICATION OF THE
LANTIBIOTIC EPIDERMIN.
CC -!- SUBCELLULAR LOCATION: POSSIBLY ASSOCIATED WITH, AND ANCHORED TO,
THE CYTOPLASMIC SIDE OF THE MEMBRANE.
CC -!- SIMILARITY: TO B.SUBTILIS SPAB AND L.LACTIS NISB.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X62386; CAA44253.1; -
DR Plasmid.
KW SEQUENCE 986 AA; 116708 MW; 343CE98526847850 CRC64;

Query Match 2.9%; Score 123; DB 1; Length 986;
Best Local Similarity 18.2%; Pred. No. 1;
Matches 153; Conservative 120; Mismatches 258; Indels 310; Gaps 40;

QY 1 MSSSSWLLLSIVATAAQSTIEQAKTFLDKFHE-----AEDL--FYQSSLASWNY 50
DB 320 ISEAAYILWLLSPNHFGTKTIRNYHEFFMDKYGFQQLVNLKQLLSDINGFGPKKDSYF 379
QY 51 NNTNTEENVONMAGDKSAFLKEQSTLAQMYPLQEQNTVRLQALQOQSSVLSL 110
DB 380 SNNI-----AFLKEK-----YLLATONNSHIEIT 404
QY 111 DSKRLNTILMTSTYTGKVCNPDNPQECILLLEPGLNETMANSLDYNERLAWESWRS 170
DB 405 NDVNLEK--NTVSKINA-----PVSTEI-----YSEIYF----- 433
QY 171 EVGKQLRPLYEYVVLKNEARAHYEDYDGRGVEVNGVDYDSRGQLEDVEHTF 230
DB 434 --GNSIKG-YEDFAVISPIILGSFNAGATGFR-TGFENIK-----KKQLQKEIVH-- 480
QY 231 EEIKPLYEHLHAYVRAKLMNAYPSYISPGCLPAHLGLDMKGRFTWNLYSILVFPFGOKPN 290
DB 481 -----HYANNMNGLEISQLNEGLSRNVLNIN--NRIYNTCLNINLP---KSD 527
QY 291 IDVTDAWQDAQRIKFKE---AEKFFSVGLPNNMTQG-----FWENSMLTDPG 337
DB 528 IDINDIFGATFNKLYLYSEKHSRIYFVSNSMFNFEFGSELYKFLREISPEKTFIOPI 587
QY 338 N-----VOKAVCHPTAWDLGKGRFRLMCKTKVTMDDELTAHHEMCHIQYDM 383
DB 588 TEEGIDSLPFCPRIYIKNIILKPAWKINSEMFSE---TENWLNRFATI-REKWHIPKDV 643

```
QY 384 AYA-----AQPELLRNGANEGFHEAVGIMS----- 409
Db 644 IIAFGDRLLLNLLDKHLILKELKKGRIILLESFINESNNRMLIIVPLPKKTSL 703
QY 410 --LSAATPK-----HLK---SIGLLSPDFQEDNETEINFLKQALTIVGLPFTYMLE 457
Db 704 KEQSFIPKRNKRNKHNKLDWFSIHLSIKPTYQD-----NFIQDVLIPTELKVNFFIN 758
QY 458 KNEWMVKEIPKDOAKMKWEMKRIVGVPVPHDETYCDPASLFHVSNDYSFIYY- 516
Db 759 KFYIKFKED--EDFIK--LRLLE-----DEDY-----SOIYFIKMKK 794
QY 517 -----TRTYQFQFQALCOAAKHEGPLHKDISNSTEAGQKLFLNMLRGKSEPWTLE 571
Db 795 DYCLLNSELYDSDVYVPEVRYGPHVIEDIENFFMYDSL--SINIQSE-FKIPKE 851
QY 572 NVYGAKMNVRLNLYFELFLTWLKDQKNSFVGWSTWSPYADOSIKVRISLKSALGDK 631
Db 852 FIVA-----ISIDFLDYLE-----INKSEKEILINNA--ED 882
QY 632 AYEWNDNEMYLFRSSVAYAMROYFLKVKQMILFGEEDVRVANLKPRIISFVFTAPKNV 691
Db 883 LYRSND-----IREY-----KNLLAKL-----TNPKNND 905
QY 692 SDIIP-----RTEVEKAIRMSRSRINDAF---RLNDSNLEFLGIQ 728
Db 906 YEILKKEFPNLHFLFNKISILENLKLTQKLSYTSRSRIISFIHMRGN-----RIFGIN 961
QY 729 P 729
Db 962 P 962

RESULT 14
BGAL_BACME STANDARD; PRT; 1034 AA.
AC 052847;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE Beta-galactosidase (EC 3.2.1.23) (Lactase).
GN BGAM.
OS Bacillus megaterium.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1404;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 319;
RA Strey J.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-
CC galactose residues in beta-D-galactosides.
CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF GLYCOSYL HYDROLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AJ000733; CAA04267.1;
CC HSP: P00722; IBSL.
CC InterPro: IPR004200; Bgal_small_C.
CC InterPro: IPR004199; Bgal_small_N.
CC InterPro: IPR001649; Glyco_hydro_2.
CC Pfam: PF02930; Bgal_small_C; 1.
CC Pfam: PF02930; Bgal_small_N; 1.
CC Pfam: PF00703; Glyco_hydro_2; 1.
CC Pfam: PF02836; Glyco_hydro_2_C; 1.
```

```
DR Pfam: PF02837; Glyco_hydro_2_N; 1.
DR PRINTS: PR00132; GLHYDRASE2.
DR PROSITE: PS00719; GLYCOSYL_HYDROL_F2_1; 1.
DR PROSITE: PS00608; GLYCOSYL_HYDROL_F2_2; 1.
KW Hydrolase; Glycosidase.
FT ACT_SITE 481 PROTON DONOR (BY SIMILARITY).
FT ACT_SITE 547 NUCLEOPHILE (BY SIMILARITY).
SQ SEQUENCE 1034 AA; 118673 MW; 38644C9A649415E9 CRC64;

Query Match 2.8%; Score 121; DB 1; Length 1034;
Best Local Similarity 19.2%; Pred. No. 1.5;
Matches 104; Conservative 66; Mismatches 165; Indels 206; Gaps 30;

QY 29 LDKFNHEAEDLFYQSSL-----ASWNYNTNITEENVQNMN---NAGDKKSAFLKE 75
Db 260 LNYFEDHQDITFVFNLYDANAQEVLPLOTNLSVQDRTVSLRTHIKSPAKWSA----- 315
QY 76 QSTLAQMYPLQEIQNL-TVKIQALQALQONGSSVLSEDKSK---RLNTILNTMTSYTSGK 131
Db 316 -----ESPPLYTLVL---SLKNAAGSIITETESCKVGFRTEIKNGLMI--NGK 359
QY 132 -----VCNPDNPQECILLPEPLNEIMANSLDYNERLW----- 163
Db 360 RIVLRGNVRHFEFSYKGRAGITREDMIHDILLMKOHINAVRTSHYPNDSVWYELCNEYG 419
QY 164 -----AW-----ESWRSEVKGKOLRPLYEE----- 182
Db 420 LYVIDETNLEHTGWTYLYQEGEKAVPGSKPEWKENVLDRCRMYERDKNHPISIIISLG 479
QY 183 -----YVVLK-NEMARANHYEDGYWRGDEVNGVDGYDYSRQGLIEDVEHT 229
Db 480 NESFGGENFQHMVTFEKKEDSTRLVHYE--GIFHHRDYDASDIESTMYVKP---ADVE-- 532
QY 230 FEEIKPLYEHLHAVRAKLMNAYSITSPICLPAHLGLDMGRFWMNLVSLVYPPGQKP 289
Db 533 -----RYALMNPKKPYIL---CEYSHAMGNSCG---NLRYKWEFLFDQYP 570
QY 290 NID---VTDMVDAQDAQRIFKEAE--KFFVSVG-----LPNMTQGFWENSMLTDPGNV 339
Db 571 ILQGGFIWD-----WKDQALQATAEADGTSYLAJGGFGDTPN-DGNFCGGLIFADGTA 623
QY 340 QKAV-----CH-PTAW---DLGKGFRLIMCTKYVTMDDFLTAHHEMGHI-----QYDMAY 385
Db 624 SPKIAEVKKCYQPVKWTAVDPAKGFV-----QNKHLFTNLNAYDFVW 667
QY 386 AAQPFLLRNGANEGFHEAVGIMSLSRAATPKHLSIGLLSPDFQEDNETEINFLKQALT 445
Db 668 TVE-----KNG-----ELVEKHASLLNVPADGTDLTSLYPLQEENETD-EFVLTLSLR 716
QY 446 I 446
Db 717 L 717

RESULT 15
YAMB_SCHPO STANDARD; PRT; 3655 AA.
ID YAMB_SCHPO
AC Q10064;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 420.8 kDa protein Clf5.11C in chromosome I.
GN SPAC1F5.11C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Gentles S.; Odell C.; Churcher C.M.; Barrell B.G.; Rajandream M.A.;
RA Walsh S.V.;
```

```
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: STRONG, TO YEAST YHR099W.
CC -!- SIMILARITY: BELONGS TO THE PI3/PI4-KINASES FAMILY
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z68136; CRA92239.1; -
DR InterPro; IPR003151; FATC.
DR InterPro; IPR003152; FATC.
DR InterPro; IPR000403; PI3_PI4_kinase.
DR InterPro; IPR001440; TFR.
DR Pfam; PF02259; FAT; 1.
DR Pfam; PF02260; FATC; 1.
DR Pfam; PF00454; PI3_PI4_kinase; 1.
DR SMART; SM00146; PI3Kc; 1.
DR PROSITE; PS50290; PI3_4_KINASE_3; 1.
KW Hypothetical protein; Transferase; Kinase.
FT DOMAIN 3324 3655 PI3K/PI4K (BY SIMILARITY).
SQ SEQUENCE 3655 AA; 420774 MW; 50475E3FC2124A CRC64;

Query Match          2.8%; Score 119.5; DB 1; Length 3655;
Best Local Similarity 17.0%; Pred. No. 11;
Matches 93; Conservative 93; Mismatches 197; Indels 165; Gaps 23;

QY 4 SSWLLLSLVAVTAAGSTIEEQAKT---FLDKFNHEADLFY-----QSSIASNNYNTNITE 56
DB 2633 SDW-----STEQDTLEKATKSLSPETSLRRHTADALLYLNKTRKMGSVTEFSRIID 2684

QY 57 ENVQNNNAGDKWSAFLE--QSTLAQMPLOEIQNLTKLQLOALQONGSSVLSEDKSK 114
DB 2685 ECMQFSLR---RWQQLPRKVYSHVSLHHEFOEIVEL-----QEAFCIYSQLNDT 2731

QY 115 RLNTILNTMTSYSGKVCNPDNPQCEILLEPGLNEIMANSLDYNERLWA-WESWRSEVG 173
DB 2732 NIHNIDNKL-----RDIKVLQGWRELRPNWD-DIDIWSDLIARQSVF 2775

QY 174 KQLRPLYEYVVLKEMARANYEDYGDYWRGDYEVNGVDGYDSRGQLIEDVEHTPEEI 233
DB 2776 KSINKVFLPLVSIQAQSTNKSNTNSVLYRGYHEL----- 2811

QY 234 KPLYEHLHAYVRAKLMNAYPSVISPIGCLPAHLGLDMGCRFWTNLYSLTPPGQKPNIDV 293
DB 2812 -----AWIINRF-AHVARVHHLPEVCINQL-----TKIYTL-----PNIEI 2846

QY 294 TDAMVDQAMDQRIFKEAEKFFVSVGLPNMT-----QGFWENSMLTDPG 337
DB 2847 QEAFKLREQAECHYESPSEMQLGLEVINNTNLMYFRNROKAEFFTLKGMFON-RLGEKD 2905

QY 338 NVOKAVCHPTANDLKGDPRIILMCKTKVTMDDFLTAHHEMGHIQYDMAYAAQPFLLRNGAN 397
DB 2906 EANAQAFATAVQIDIGSG-----KAWSEWG-LYHDELFOANPQEIHHACN 2948

QY 398 --EGFHEAVGEIMSLSAATPKHLKLSIGLLSPDFQEDNETEINFLKQALTIVGTLPFTYM 455
DB 2949 AVSCFLQA--SSLLSSNSKPLTRVNLWLLSVDDSDSGSSEV-----VSSFKSE 2995

QY 456 LEKRWNVFKGEIPQDMKKWEMKREIVGVVEPVPHDetyCDPASLPHVSNdy-SFIR 514
DB 2996 IPTWNWIPF---IPQ-----LLSALSHRSHRAILIQIAKTYPSLIH 3036

QY 515 YVTRFLYQ 522
DB 3037 FQLRTAYE 3044
```